

Connection between Annealed Free Energy and Belief Propagation on Random Factor Graph Ensembles

Ryuhei Mori

Graduate School of Informatics, Kyoto University
Kyoto, 606-8501, Japan
Email: rmori@sys.i.kyoto-u.ac.jp

Abstract—Recently, Vontobel showed the relationship between Bethe free energy and annealed free energy for protograph factor graph ensembles. In this paper, annealed free energy of any random regular, irregular and Poisson factor graph ensembles are connected to Bethe free energy. The annealed free energy is expressed as the solution of maximization problem whose stationary condition equations coincide with equations of belief propagation since the contribution to partition function of particular type of variable and factor nodes has similar form of minus Bethe free energy. It gives simple derivation of replica symmetric solution. As consequence, it is shown that on replica symmetric ansatz, replica symmetric solution and annealed free energy are equal for regular ensemble.

I. INTRODUCTION

In the context of statistical physics, free energy of disordered system is central interest. In information theory, the a posteriori distribution of low-density parity-check (LDPC) codes can be regarded as Boltzmann-Gibbs distributions on sparse factor graphs whose free energy is related to the conditional entropy of codewords under a received vector [1]. In computer science, constraint satisfaction problems (CSPs) which can be expressed by sparse factor graphs are important theoretical objects. Relation between phase transition phenomenon and free energy of randomized CSPs has also been considered well [2], [3].

In this paper, we deal with calculation of *annealed free energy* of random sparse factor graph ensemble on finite alphabet. Although in many cases *quenched free energy* gives meaningful result e.g., conditional entropy of LDPC codes [1], phase transition point of random CSPs [2], the calculation of quenched free energy is often difficult without *replica method* which is mathematically nonrigorous but powerful tool of statistical physics. Annealed free energy is also important quantity since it can be used for bound of quenched free energy and is required in the replica method.

For many cases [4], in the calculation of annealed and quenched free energy, fixed point equations of belief propagation (BP) and its density evolution (DE) appear, respectively. However, the relationship between BP (DE) and annealed (quenched) free energy has not been well understood. Recently, Vontobel show the relationship between Bethe free energy of protograph ensemble and its annealed free energy [5]. From this result, we can connect BP and annealed free energy

since BP equation is equivalent to stationary condition of Bethe free energy [6].

The main result of this paper is derivation of annealed free energy of any random regular, irregular and Poisson factor graph ensembles by using BP equations. The derivation of annealed free energy gives the simple derivation of replica symmetric solution. It is shown that if the replica symmetric assumption is correct, annealed and quenched free energy are equal for any regular ensembles.

II. FACTOR GRAPH, GIBBS FREE ENERGY AND BETHE APPROXIMATION

In this paper, we deal with factor graph which is bipartite graph representing probability distribution [6], [3]. Let us consider bipartite graph consists of N variable nodes and M factor nodes. Let \mathcal{X} be alphabet which is common domain of variables. For each factor node a , there is a function $f_a : \mathcal{X}^{r_a} \rightarrow \mathbb{R}_{\geq 0}$ where r_a denotes the degree of a . The factor graph represents the following distribution p on \mathcal{X}^N .

$$p(\mathbf{x}) = \frac{1}{Z} \prod_a f_a(\mathbf{x}_{\partial a})$$

where

$$Z := \sum_{\mathbf{x}} \prod_a f_a(\mathbf{x}_{\partial a})$$

is constant for normalization, i.e., $\sum_{\mathbf{x}} p(\mathbf{x}) = 1$. Here, $\mathbf{x}_{\partial a}$ denotes value of variable nodes connecting a factor node a . In the context of statistical mechanics, Z is called partition function and $-\log Z$ is called Helmholtz free energy.

When N is large, calculation of Z requires large computational complexity. Hence, the approximation of p by simple distribution q is often introduced. The following method of approximation is written in [6]. For the criteria of approximation, Kullback-Leibler divergence is used.

$$\begin{aligned} D(q||p) &:= \sum_{\mathbf{x}} q(\mathbf{x}) \log \frac{q(\mathbf{x})}{p(\mathbf{x})} \\ &= \log Z - \sum_{\mathbf{x}} \sum_a q(\mathbf{x}) \log f_a(\mathbf{x}_{\partial a}) + \sum_{\mathbf{x}} q(\mathbf{x}) \log q(\mathbf{x}) \\ &=: \log Z + \mathcal{U}(q) - \mathcal{H}(q) =: \log Z + \mathcal{F}_{\text{Gibbs}}(q) \end{aligned}$$

The quantity $\mathcal{U}(q)$, $\mathcal{H}(q)$ and $\mathcal{F}_{\text{Gibbs}}(q)$ are called internal energy, entropy and Gibbs free energy, respectively.

The approximation using $q(\mathbf{x})$ which is factorized as $\prod_{i=1}^N q_i(x_i)$, i.e., x_i are independent, is called mean field approximation. The approximation using $q(\mathbf{x})$ which is represented as

$$q(\mathbf{x}) = \frac{\prod_a b_a(\mathbf{x}_{\partial a})}{\prod_i b_i(x_i)^{l_i-1}}$$

is called *Bethe approximation* where i and a represent indices of variable nodes and factor nodes, respectively, and where l_i denotes degree of variable node i . For Bethe approximation, Bethe average energy and Bethe entropy are defined as

$$\begin{aligned} \mathcal{U}_{\text{Bethe}}(b_a) &:= - \sum_a \sum_{\mathbf{x}_{\partial a}} b_a(\mathbf{x}_{\partial a}) \log f_a(\mathbf{x}_{\partial a}) \\ \mathcal{H}_{\text{Bethe}}(b_i, b_a) &:= - \sum_a \sum_{\mathbf{x}_{\partial a}} b_a(\mathbf{x}_{\partial a}) \log b_a(\mathbf{x}_{\partial a}) \\ &\quad + \sum_i \sum_{x_i} (l_i - 1) b_i(x_i) \log b_i(x_i) \end{aligned} \quad (1)$$

respectively. Bethe free energy is defined as $\mathcal{F}_{\text{Bethe}}(b_i, b_a) := \mathcal{U}_{\text{Bethe}}(b_a) - \mathcal{H}_{\text{Bethe}}(b_i, b_a)$. In order to obtain good Bethe approximation, minimization of Bethe free energy is considered since Bethe free energy is analogy of Gibbs free energy, whose minimization is equivalent to minimization of the Kullback-Leibler divergence. When we assume constraints, $\sum_i b_i(x_i) = 1$ for all variable nodes i , $\sum_{\mathbf{x}_{\partial a}} b_a(\mathbf{x}_{\partial a}) = 1$ for all factor nodes a , and $\sum_{\mathbf{x}_{\partial a}, x_i=x} b_a(\mathbf{x}_{\partial a}) = b_i(x)$ for all factor nodes a and variable nodes $i \in \partial a$, the stationary condition of Lagrangian is equivalent to condition of fixed point of BP [6].

III. ANNEALED FREE ENERGY OF RANDOM REGULAR FACTOR GRAPH ENSEMBLES

In this paper, we mainly deal with random regular factor graph ensembles. Results for regular ensembles can be generalized straightforwardly to irregular and Poisson ensembles. Let l and r be degrees of variable and factor nodes of regular factor graph ensembles, respectively. Let $\mathbb{E}[\cdot]$ denote the expectation on random connection of edges. Two quantities $\mathbb{E}[\log Z]$ and $\log \mathbb{E}[Z]$ are called quenched and annealed free energy, respectively. The main purpose of this paper is calculation of $\lim_{N \rightarrow \infty} 1/N \log \mathbb{E}[Z]$ where N denotes the number of variable nodes. The essential idea of calculation is type classification of the contribution to partition function [5]. Let *variable-type* v denote the type of variable nodes, i.e., there exists $v(x)$ variable nodes of value $x \in \mathcal{X}$. Let *factor-type* u denote the type of factor nodes, in which the value of factor nodes is regarded as the values of variable nodes connects to the factor nodes, i.e., there exists $u(\mathbf{x})$ factor nodes connecting variable nodes of value $\mathbf{x} \in \mathcal{X}^r$. In this paper, for simplicity, factors $f_a(\mathbf{x}_{\partial a})$ do not depend on factor node a , and written as $f(\mathbf{x}_{\partial a})$. Let $Z(v, u)$ be the contribution of assignments with variable-type v and factor-type u , and $N(v, u)$ be the number of assignments with variable-type v and factor-type u .

$$Z = \sum_{v, u} Z(v, u) = \sum_{v, u} N(v, u) \prod_{\mathbf{x} \in \mathcal{X}^r} f(\mathbf{x})^{u(\mathbf{x})}$$

In the sum, the types v and u have to satisfy the consistency condition

$$\sum_{i=1}^r \sum_{\substack{\mathbf{x} \setminus x_i \\ x_i=z}} u(\mathbf{x}) = lv(z).$$

The number $N(v, u)$ of assignments with variable-type v and factor-type u is

$$\mathbb{E}[N(v, u)] = \binom{N}{\{v(x)\}_{x \in \mathcal{X}}} \binom{\frac{l}{r}N}{\{u(\mathbf{x})\}_{\mathbf{x} \in \mathcal{X}^r}} \frac{\prod_{x \in \mathcal{X}} (v(x)l)!}{(Nl)!}.$$

Now, we consider the exponent of the contribution of types v and u where $v(x) := v(x)/N$ and $\mu(\mathbf{x}) := u(\mathbf{x})/((l/r)N)$, respectively. It holds

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}[Z(v, \mu)] \\ = \frac{l}{r} \mathcal{H}(\mu) - (l-1) \mathcal{H}(v) + \frac{l}{r} \sum_{\mathbf{x} \in \mathcal{X}^r} \mu(\mathbf{x}) \log f(\mathbf{x}) \\ =: -F_{\text{Bethe}}(v, \mu). \end{aligned}$$

Hence,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}[Z] = \max_{v, \mu} \{-F_{\text{Bethe}}(v, \mu)\}$$

where, v and μ have to satisfy the following conditions.

$$\begin{aligned} v(x) &\geq 0, \forall x \in \mathcal{X}, & \mu(\mathbf{x}) &\geq 0, \forall \mathbf{x} \in \mathcal{X}^r \\ \sum_{x \in \mathcal{X}} v(x) &= 1, & \sum_{\mathbf{x} \in \mathcal{X}^r} \mu(\mathbf{x}) &= 1, \end{aligned}$$

$$\frac{1}{r} \sum_{i=1}^r \sum_{\substack{\mathbf{x} \setminus x_i \\ x_i=z}} \mu(\mathbf{x}) = v(z), \forall z \in \mathcal{X}.$$

The last condition is for the consistency between v and μ . The above maximization problem is similar to the minimization problem of Bethe free energy. Hence, we can easily understand that the stationary condition is similar to the fixed point equation of BP. The Lagrangian of the maximization problem is

$$\begin{aligned} L(v, \mu; \lambda, \rho, \tau) &= -F_{\text{Bethe}}(v, \mu) \\ &\quad + \lambda \left(\sum_{x \in \mathcal{X}} v(x) - 1 \right) + \frac{l}{r} \rho \left(\sum_{\mathbf{x} \in \mathcal{X}^r} \mu(\mathbf{x}) - 1 \right) \\ &\quad + \sum_{z \in \mathcal{X}} \tau(z) \left(\frac{l}{r} \sum_{i=1}^r \sum_{\substack{\mathbf{x} \setminus x_i \\ x_i=z}} \mu(\mathbf{x}) - lv(z) \right). \end{aligned} \quad (2)$$

Lemma 1. The stationary condition of (2) is

$$\begin{aligned} v(x) &\propto m_{f \rightarrow v}(x)^l \\ \mu(\mathbf{x}) &\propto f(\mathbf{x}) \prod_{i=1}^r m_{v \rightarrow f}(x_i) \end{aligned}$$

where

$$m_{v \rightarrow f}(x) \propto m_{f \rightarrow v}(x)^{l-1} \quad (3)$$

$$m_{f \rightarrow v}(x) \propto \sum_{i=1}^r \sum_{\substack{\mathbf{x} \setminus x_i \\ x_i=x}} f(\mathbf{x}) \prod_{j=1, j \neq i}^r m_{v \rightarrow f}(x_j). \quad (4)$$

Here $m_{v \rightarrow f}(x)$ and $m_{f \rightarrow v}(x)$ are auxiliary functions satisfying $\sum_{x \in \mathcal{X}} m_{v \rightarrow f}(x) = \sum_{x \in \mathcal{X}} m_{f \rightarrow v}(x) = 1$.

Proof is in Appendix A. If $f(\mathbf{x})$ is invariant under permutation of \mathbf{x} , (4) is simply written as

$$m_{f \rightarrow v}(x) \propto \sum_{\substack{\mathbf{x} \setminus x_1 \\ x_1 = x}} f(\mathbf{x}) \prod_{j=2}^r m_{v \rightarrow f}(x_j).$$

Theorem 2.

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}[Z] = \max_{(m_{v \rightarrow f}(x), m_{f \rightarrow v}(x)) \in \mathcal{S}} \left\{ \frac{l}{r} \log Z_f + \log Z_v - l \log Z_{fv} \right\}.$$

where \mathcal{S} denotes the set of saddle points of the function in max, and where

$$\begin{aligned} Z_v &:= \sum_x m_{f \rightarrow v}(x)^l \\ Z_f &:= \sum_{\mathbf{x}} f(\mathbf{x}) \prod_{i=1}^r m_{v \rightarrow f}(x_i) \\ Z_{fv} &:= \sum_x m_{f \rightarrow v}(x) m_{v \rightarrow f}(x). \end{aligned}$$

The conditions of saddle point are (3) and (4).

Proof is in Appendix B.

Remark 3. Assume that $\sum_{i=1}^r \sum_{\substack{\mathbf{x} \setminus x_i \\ x_i = x}} f(\mathbf{x})$ is constant among all $x \in \mathcal{X}$. Then, the uniform distributions $m_{v \rightarrow f}(x)$ and $m_{f \rightarrow v}(x)$ are a trivial fixed point. Let $N_f := \sum_{\mathbf{x}} f(\mathbf{x})$. The contribution $Z(v, \mu)$ evaluated at uniform v and μ is

$$\frac{l}{r} \log \frac{N_f}{q^r} + \log \frac{1}{q^{l-1}} - l \log \frac{1}{q} = \log q + \frac{l}{r} \log \frac{N_f}{q^r}. \quad (5)$$

When $f(\mathbf{x}) \in \{0, 1\}$, i.e., the problem is the CSP, Z is the number of solution and N_f is the cardinality of $\{\mathbf{x} \in \mathcal{X}^r \mid f(\mathbf{x}) = 1\}$. In this case, we call the quantity (5) *design rate*. If the uniform v and μ maximize $Z(v, \mu)$, the expected number $\mathbb{E}[Z]$ of solution is about

$$q^N \left(\frac{N_f}{q^r} \right)^{\frac{l}{r}}.$$

Roughly speaking, this implies that all constraints are independent. This solution is called *paramagnetic solution* in [2], in the context of replica symmetric solution.

The generalization for irregular and Poisson ensemble is in Appendix H.

IV. CONTRIBUTION TO PARTITION FUNCTION OF FIXED VARIABLE TYPE

We now consider the contribution to partition function of regular factor graph ensemble with fixed variable type. More precisely, we consider $Z(v) := \sum_u Z(v, u)$. It holds

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}[Z(v)] = \max_{\mu} \{-F_{\text{Bethe}}(v, \mu)\}. \quad (6)$$

The function $-F_{\text{Bethe}}(v, \mu)$ is a concave function with respect to μ . Since the equality constraints are linear, the problem

is essentially a maximization problem of a concave function without constraints. Hence, it can be solved numerically by the Newton method.

Lemma 4. The stationary condition of (6) is

$$\mu(\mathbf{x}) \propto f(\mathbf{x}) \prod_{i=1}^r m_{v \rightarrow f}(x_i)$$

where

$$v(x) \propto h(x) m_{f \rightarrow v}(x)^l \quad (7)$$

$$m_{v \rightarrow f}(x) \propto h(x) m_{f \rightarrow v}(x)^{l-1} \quad (8)$$

$$m_{f \rightarrow v}(x) \propto \sum_{i=1}^r \sum_{\substack{\mathbf{x} \setminus x_i \\ x_i = x}} f(\mathbf{x}) \prod_{j=1, j \neq i}^r m_{v \rightarrow f}(x_j). \quad (9)$$

Here $m_{v \rightarrow f}(x)$, $m_{f \rightarrow v}(x)$ and $h(x)$ are auxiliary functions satisfying $\sum_{x \in \mathcal{X}} m_{v \rightarrow f}(x) = \sum_{x \in \mathcal{X}} m_{f \rightarrow v}(x) = 1$, and $h(x) \geq 0$.

Proof is in Appendix D. Since $h(x)$ is arbitrary auxiliary function, $m_{f \rightarrow v}(x)^l$ in (7) and $m_{f \rightarrow v}(x)^{l-1}$ in (8), can be replaced by $m_{f \rightarrow v}(x)^k$ and $m_{f \rightarrow v}(x)^{k-1}$, respectively for any $k \geq 1$. Here, we chose $k = l$ since we can obtain the following simple result. The stationary condition for magnetic field model in Appendix C and Lemma 4 are the similar although while in the problem for magnetic field, $h(x)$ is given and $v(x)$ is variable, in this problem, $h(x)$ is variable and $v(x)$ is given.

Lemma 5.

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}[Z(v)] &= \max_{(m_{f \rightarrow v}(x), m_{v \rightarrow f}(x), h(x)) \in \mathcal{S}} \left\{ \frac{l}{r} \log Z_f + \log Z_v - l \log Z_{fv} \right. \\ &\quad \left. - \sum_x v(x) \log h(x) \right\} \\ &= \max_{(m_{f \rightarrow v}(x), m_{v \rightarrow f}(x)) \in \mathcal{S}} \left\{ \frac{l}{r} \log Z_f + \sum_x v(x) \log Z_v(x) - l \log Z_{fv} \right\} \\ &\quad + \mathcal{H}(v). \end{aligned} \quad (10)$$

where \mathcal{S} denotes the set of saddle points of the function in max, and where

$$\begin{aligned} Z_f &:= \sum_{\mathbf{x}} f(\mathbf{x}) \prod_{i=1}^r m_{v \rightarrow f}(x_i) \\ Z_v(x) &:= m_{f \rightarrow v}(x)^l \\ Z_v &:= \sum_x h(x) Z_v(x) \\ Z_{fv} &:= \sum_x m_{f \rightarrow v}(x) m_{v \rightarrow f}(x). \end{aligned}$$

The conditions of saddle point are (7), (8) and (9).

The annealed free energy of magnetic field model in Appendix C is obtained by the Legendre transform of the above result. It can be easily verified from (10).

While the maximization problem (6) can be solved by the Newton method, Lemma 4 gives the efficient algorithm. First,

$\{m_{f \rightarrow v}^{(0)}(x)\}_{x \in \mathcal{X}}$ are initialized. Then, messages are updated by

$$m_{v \rightarrow f}^{(t+1)}(x) \propto \frac{v(x)}{m_{f \rightarrow v}^{(t)}(x)}$$

$$m_{f \rightarrow v}^{(t)}(x) \propto \sum_{i=1}^r \sum_{\substack{\mathbf{x}_{\setminus x_i} \\ x_i = x}} f(\mathbf{x}) \prod_{j=1, j \neq i}^r m_{v \rightarrow f}^{(t)}(x_j)$$

iteratively. After sufficient iterations, messages are substituted to

$$l \left(\frac{1}{r} \log Z_f + \sum_x v(x) \log m_{f \rightarrow v}(x) - \log Z_{fv} \right) + H(v).$$

Note that the degree l of variable nodes does not appear in the iterations and only appear as the factor of the first term in the last equation. This algorithm does not necessarily converges. The example of problem for which the above BP-like algorithm does not converge is shown in Section VI.

V. MOMENT OF PARTITION FUNCTION AND REPLICA METHOD

In this section, we deal with moments of partition function which is useful for some purposes. One of the most successful result of use of moment is the second moment method i.e., for nonnegative random variable Z , $P(Z > 0) \geq \mathbb{E}[Z]^2 / \mathbb{E}[Z^2]$. Using this method, lower bound of SAT-UNSAT threshold is obtained [7]. The other use of moment is the replica method which is not rigorous but powerful tool of statistical physics for calculation of quenched free energy. The basic idea of the replica method is representation of $\mathbb{E}[\log Z]$ as the derivative $(\partial \log \mathbb{E}[Z^n]) / \partial n|_{n=0}$. It holds

$$\lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}[\log Z] = \lim_{N \rightarrow \infty} \frac{1}{N} \lim_{n \rightarrow 0} \frac{\log \mathbb{E}[Z^n]}{n}.$$

If the exchange of the limits is admissible,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}[\log Z] = \lim_{n \rightarrow 0} \frac{1}{n} \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}[Z^n]. \quad (11)$$

In the replica method,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}[Z^n] \quad (12)$$

have to be evaluated. Usually, (12) is evaluated only for $n \in \mathbb{N}$ such that dependence on n is analytic. Then, the right-hand side of (11) is evaluated by ignoring that n should be natural number [3].

Since Z^n can be regarded as partition function of factor graph on alphabet \mathcal{X}^n and factor $\prod_{i=1}^n f(\mathbf{x}^{(i)})$, the exponent of moment is also calculated in the same way. Here, $\mathbf{x}^{(i)} \in \mathcal{X}^r$ denotes vector $(\mathbf{x}_1^{(i)}, \dots, \mathbf{x}_r^{(i)})$ where \mathbf{x}_j is j -th elements of $\mathbf{x} \in (\mathcal{X}^n)^r$ and $\mathbf{x}_j^{(i)}$ denotes i -th element of $\mathbf{x}_j \in \mathcal{X}^n$.

Corollary 6.

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}[Z^n] = \max_{(m_{f \rightarrow v}(\mathbf{x}), m_{v \rightarrow f}(\mathbf{x})) \in \mathcal{S}} \left\{ \frac{l}{r} \log Z_f + \log Z_v - l \log Z_{fv} \right\} \quad (13)$$

where \mathcal{S} denotes the set of saddle points of the function in max, and where

$$Z_v := \sum_{\mathbf{x} \in \mathcal{X}^n} m_{f \rightarrow v}(\mathbf{x})^l$$

$$Z_f := \sum_{\mathbf{x} \in (\mathcal{X}^n)^r} \left(\prod_{j=1}^n f(\mathbf{x}^{(j)}) \right) \prod_{i=1}^r m_{v \rightarrow f}(\mathbf{x}_i)$$

$$Z_{fv} := \sum_{\mathbf{x} \in \mathcal{X}^n} m_{f \rightarrow v}(\mathbf{x}) m_{v \rightarrow f}(\mathbf{x}).$$

The essentially same result for LDPC codes was obtained in [8] (Eq. (5.2)). In [8], it is explained that the replica symmetric assumption says that distributions $m_{v \rightarrow f}(x^{(1)}, \dots, x^{(n)})$ and $m_{f \rightarrow v}(x^{(1)}, \dots, x^{(n)})$ which are invariant under permutation dominates $\mathbb{E}[Z^n]$. Furthermore, the representations

$$m_{v \rightarrow f}(\mathbf{x}) = \int \prod_{i=1}^n M_{v \rightarrow f}(x_i) d\Phi(M_{v \rightarrow f})$$

$$m_{f \rightarrow v}(\mathbf{x}) = \int \prod_{i=1}^n M_{f \rightarrow v}(x_i) d\hat{\Phi}(M_{f \rightarrow v})$$

are assumed where Φ and $\hat{\Phi}$ denote probability measures on $\mathcal{P}(\mathcal{X})$, i.e., Φ and $\hat{\Phi}$ are elements of $\mathcal{P}(\mathcal{P}(\mathcal{X}))$. Here, $\mathcal{P}(\mathcal{A})$ denotes the set of probability measures on a set \mathcal{A} .

Lemma 7.

$$-F_{RS} = \max_{(\Phi, \hat{\Phi}) \in \mathcal{S}} \left\{ \frac{l}{r} \langle \log Z_f \rangle + \langle \log Z_v \rangle - l \langle \log Z_{fv} \rangle \right\}$$

where \mathcal{S} denotes the set of saddle points of the function in max, where

$$Z_v := \sum_{\mathbf{x} \in \mathcal{X}^n} \prod_{i=1}^l M_{f \rightarrow v}^{(i)}(x)$$

$$Z_f := \sum_{\mathbf{x} \in \mathcal{X}^r} f(\mathbf{x}) \prod_{i=1}^r M_{v \rightarrow f}^{(i)}(x_i)$$

$$Z_{fv} := \sum_{\mathbf{x} \in \mathcal{X}} M_{v \rightarrow f}(x) M_{f \rightarrow v}(x)$$

where $\{M_{v \rightarrow f}^{(i)}\}_{i=1, \dots, r}$ and $\{M_{f \rightarrow v}^{(i)}\}_{i=1, \dots, l}$ are i.i.d. random measures obeying Φ and $\hat{\Phi}$, respectively, and where $\langle \cdot \rangle$ denotes the expectation with respect to the random measures. The saddle point conditions are

$$\frac{\prod_{i=1}^{l-1} M_{f \rightarrow v}^{(i)}(x)}{\sum_{\mathbf{x} \in \mathcal{X}} \prod_{i=1}^{l-1} M_{f \rightarrow v}^{(i)}(x)} \sim \Phi$$

$$\frac{\sum_{\mathbf{x} \in \mathcal{X}^r, x_D = x} f(\mathbf{x}) \prod_{j=1, j \neq D}^r M_{v \rightarrow f}^{(j)}(x_j)}{\sum_{\mathbf{x} \in \mathcal{X}^r} f(\mathbf{x}) \prod_{j=1, j \neq D}^r M_{v \rightarrow f}^{(j)}(x_j)} \sim \hat{\Phi}$$

where D denotes the uniform random variable on $\{1, 2, \dots, r\}$ which is independent of any random variable, and where $M \sim \Phi$ denotes that a random measure M has a law Φ .

Proof is in Appendix E. This derivation of replica symmetric solution is simpler than previously known ones [9], [8], [10] in which complicated tools are used e.g., integral expression of delta function. Another advantage of this paper is that we

can understand why the saddle point equation in the replica symmetric solution is equal to the DE equation.

When $f(\mathbf{x})$ is invariant under permutation of \mathbf{x} , the fixed points for annealed free energy in Lemma 1 are also fixed point for RS saddle point equation as delta distribution. From inclusion relation of domains of max in Theorem 2 and Lemma 7, $-F_{RS} \geq \lim_{N \rightarrow \infty} 1/N \log \mathbb{E}[Z]$. On the other hand, from Jensen's inequality, $\mathbb{E}[\log Z] \leq \log \mathbb{E}[Z]$. We now obtain the following theorem.

Theorem 8. Assume $f(\mathbf{x})$ is invariant under permutation of \mathbf{x} . If replica symmetric assumption is valid i.e., $-F_{RS} = \lim_{N \rightarrow \infty} 1/N \mathbb{E}[\log Z]$, then $\lim_{N \rightarrow \infty} 1/N \mathbb{E}[\log Z] = \lim_{N \rightarrow \infty} 1/N \log \mathbb{E}[Z]$.

This result is well known for regular LDPC codes [10]. When we believe the replica method, even if replica symmetric assumption is not valid, intuitively $-F_{RS} \leq \lim_{N \rightarrow \infty} 1/N \mathbb{E}[\log Z]$ holds, since the replica symmetric assumption restrict the domain of maximization problem. However, generally $-F_{RS} \geq \lim_{N \rightarrow \infty} 1/N \mathbb{E}[\log Z]$ can be hold [11]. Hence, Theorem 8 requires the replica symmetric assumption.

This result can be generalized for random factor model straightforwardly. For the random magnetic field model in Appendix C, $\lim_{N \rightarrow \infty} 1/N \log \mathbb{E}_{\{h_i\}}[\mathbb{E}[Z^n]]$ have to be evaluated for the replica method. This quantity can be calculated easily by Theorem 2 by replacing $h(x)$ by $\mathbb{E}_h[h(x)]$. In this case, the relation $-F_{RS} \geq \lim_{N \rightarrow \infty} 1/N \log \mathbb{E}_{\{h_i\}}[\mathbb{E}[Z]]$ does not hold.

VI. APPLICATIONS

In this section, an example of binary CSP is shown. The factor is

$$f(\mathbf{x}) = \begin{cases} 0, & \text{if } \frac{r}{2} - k < \sum_{i=1}^r x_i < \frac{r}{2} + k \\ 1, & \text{otherwise} \end{cases} \text{ for } \mathbf{x} \in \{0, 1\}^r. \quad (14)$$

This factor is considered to prevent assignment from including half numbers of 0s and 1s. The number of solution of fixed variable type is calculated by the BP-like algorithm shown in Section IV. The calculation results for (10,20) regular ensemble are shown in Fig. 1. The horizontal axis shows the relative number of 1s in solutions. For all k , $v(1) = 1/2$ is not peak of growth rate. This means that the paramagnetic solution is not solution of the maximization problem in Theorem 2. For $k = 3$, algorithm does not converges in region including $v(1) = 1/2$. When $v(1) = 1/2$, the paramagnetic solution $m_{v \rightarrow f}(x) = m_{f \rightarrow v}(x) = 1/2$ for $x = 0, 1$ is a fixed point of the iteration. In Appendix I, the stability condition of the paramagnetic solution when $v(1) = 1/2$ is shown. It is confirmed that the stability condition is violated for $r = 20$ and $k = 3$.

VII. CONCLUSION

The annealed free energy of any regular, irregular and Poisson factor graph ensembles are shown. The expression of annealed free energy includes the BP equation. This result gives simple derivation of replica symmetric solution. As consequence, on the replica symmetric ansatz, it is shown that

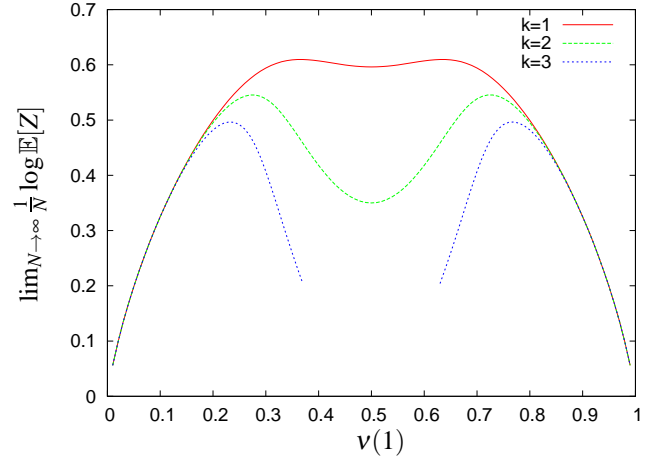


Fig. 1. Growth rate of the (10,20) ensembles.

annealed and quenched free energy are equal for any regular ensembles satisfying that $f(\mathbf{x})$ is invariant under permutation of \mathbf{x} .

ACKNOWLEDGMENT

The author acknowledges Toshiyuki Tanaka for insightful discussion. This work was supported by the Grant-in-Aid for Scientific Research for JSPS Fellows (22-5936), MEXT, Japan.

REFERENCES

- [1] N. Macris, "Griffith–Kelly–Sherman Correlation Inequalities: A Useful Tool in the Theory of Error Correcting Codes," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 664–683, 2007.
- [2] J. van Mourik and D. Saad, "Random graph coloring: Statistical physics approach," *Physical Review E*, vol. 66, no. 5, p. 56120, 2002.
- [3] M. Mezard and A. Montanari, *Information, Physics, and Computation*. Oxford University Press, USA, 2009.
- [4] C. Di, A. Montanari, and R. Urbanke, "Weight distributions of LDPC code ensembles: combinatorics meets statistical physics," in *Proc. IEEE Int. Symposium on Inform. Theory, Lausanne, Switzerland*. IEEE, 2004, p. 102.
- [5] P. Vontobel, "Counting in graph covers: A combinatorial characterization of the bethe entropy function," 2010. [Online]. Available: <http://arxiv.org/abs/1012.0065>
- [6] J. Yedidia, W. Freeman, and Y. Weiss, "Constructing free-energy approximations and generalized belief propagation algorithms," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2282–2312, 2005.
- [7] D. Achlioptas and Y. Peres, "The threshold for random k -SAT is $2k(\ln 2 - O(k))$," in *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*. ACM, 2003, pp. 223–231.
- [8] A. Montanari, "The glassy phase of Gallager codes," *The European Physical Journal B-Condensed Matter and Complex Systems*, vol. 23, no. 1, pp. 121–136, 2001.
- [9] T. Murayama, Y. Kabashima, D. Saad, and R. Vicente, "Statistical physics of regular low-density parity-check error-correcting codes," *Physical Review E*, vol. 62, no. 2, pp. 1577–1591, 2000.
- [10] S. Condamine, "Study of the weight enumerator function for a Gallager code," *Project Report Cavendish Laboratory, University of Cambridge, Cambridge, UK*, 2002.
- [11] F. Guerra, "Broken replica symmetry bounds in the mean field spin glass model," *Communications in Mathematical Physics*, vol. 233, no. 1, pp. 1–12, 2003.

APPENDIX A
PROOF OF LEMMA 1

Partial derivatives of the Lagrangian (2) are

$$\begin{aligned}\frac{\partial L}{\partial v(x)} &= (l-1)(\log v(x) + 1) + \lambda - l\tau(x) \\ \frac{\partial L}{\partial \mu(\mathbf{x})} &= -\frac{l}{r}(\log \mu(\mathbf{x}) + 1) + \frac{l}{r} \log f(\mathbf{x}) + \frac{l}{r} \rho + \frac{l}{r} \sum_{i=1}^r \tau(x_i)\end{aligned}$$

We can define $m_{v \rightarrow f}(x)$ and $m_{f \rightarrow v}(x)$ which satisfies $\sum_{x \in \mathcal{X}} m_{v \rightarrow f}(x) = \sum_{x \in \mathcal{X}} m_{f \rightarrow v}(x) = 1$ as

$$\tau(x) =: \log m_{v \rightarrow f}(x) =: \log \frac{1}{Z_{v \rightarrow f}} m_{f \rightarrow v}(x)^{l-1}$$

where $Z_{v \rightarrow f}$ is normalization constant.

$$\begin{aligned}v(x) &= \exp \left\{ -1 - \frac{\lambda}{l-1} \right\} \left(\frac{1}{Z_{v \rightarrow f}} \right)^{\frac{1}{l-1}} m_{f \rightarrow v}(x)^l \\ \mu(\mathbf{x}) &= \exp \{ -1 + \rho \} f(\mathbf{x}) \prod_{i=1}^r m_{v \rightarrow f}(x_i)\end{aligned}$$

From the normalization conditions, λ and ρ are determined uniquely. From the consistency condition of $v(x)$ and $\mu(\mathbf{x})$, it holds

$$\begin{aligned}\frac{1}{r} \sum_{i=1}^r \sum_{\mathbf{x} \setminus x_i} \frac{1}{Z_f} f(\mathbf{x}) \prod_{j=1}^r m_{v \rightarrow f}(x_j) &= \frac{1}{Z_v} m_{f \rightarrow v}(x)^l \\ \iff m_{v \rightarrow f}(x) \frac{1}{r} \sum_{i=1}^r \sum_{\mathbf{x} \setminus x_i} \frac{1}{Z_f} f(\mathbf{x}) \prod_{j=1, j \neq i}^r m_{v \rightarrow f}(x_j) &= \frac{1}{Z_v} m_{f \rightarrow v}(x)^l \\ \iff \frac{Z_v}{Z_f Z_{v \rightarrow f}} \frac{1}{r} \sum_{i=1}^r \sum_{\mathbf{x} \setminus x_i} f(\mathbf{x}) \prod_{j=1, j \neq i}^r m_{v \rightarrow f}(x_j) &= m_{f \rightarrow v}(x).\end{aligned}$$

APPENDIX B
PROOF OF THEOREM 2

Let us consider

$$\frac{l}{r} \mathcal{H}(\mu) - (l-1) \mathcal{H}(v) + \frac{l}{r} \sum_{\mathbf{x} \in \mathcal{X}^r} \mu(\mathbf{x}) \log f(\mathbf{x}) \quad (15)$$

evaluated at v and μ at the stationary point i.e., they satisfy Lemma 1.

$$\begin{aligned}\frac{l}{r} \mathcal{H}(\mu) &= \frac{l}{r} \log Z_f - \frac{l}{r} \sum_{\mathbf{x}} \mu(\mathbf{x}) \log \left(f(\mathbf{x}) \prod_{i=1}^r m_{v \rightarrow f}(x_i) \right) \\ &= \frac{l}{r} \log Z_f - \frac{l}{r} \sum_{\mathbf{x}} \mu(\mathbf{x}) \log f(\mathbf{x}) - l \sum_{\mathbf{x}} v(x) \log m_{v \rightarrow f}(x) \\ &= \frac{l}{r} \log Z_f - \frac{l}{r} \sum_{\mathbf{x}} \mu(\mathbf{x}) \log f(\mathbf{x}) - l \sum_{\mathbf{x}} v(x) \log m_{f \rightarrow v}(x)^{l-1} \\ &\quad + l \log Z_{v \rightarrow f}\end{aligned}$$

$$(l-1) \mathcal{H}(v) = (l-1) \log Z_v - (l-1) \sum_x v(x) \log m_{f \rightarrow v}(x)^l$$

Hence, (15) is

$$\frac{l}{r} \log Z_f + \log Z_v - l \log \frac{Z_v}{Z_{v \rightarrow f}}$$

The equation in Theorem 2 is obtained from the equality

$$\frac{Z_v}{Z_{v \rightarrow f}} = \frac{\sum_{x \in \mathcal{X}} m_{f \rightarrow v}(x)^l}{\sum_{z \in \mathcal{X}} m_{f \rightarrow v}(z)^{l-1}} = Z_{fv}.$$

On the other hand, let us consider the function

$$F(\{m_{v \rightarrow f}\}, \{m_{f \rightarrow v}\}) := \frac{l}{r} \log Z_f + \log Z_v - l \log Z_{fv}$$

for any non-negative functions $m_{f \rightarrow v}(x)$ and $m_{v \rightarrow f}(x)$. This quantity is invariant under scaling of $\{m_{f \rightarrow v}(x)\}_{x \in \mathcal{X}}$ and $\{m_{v \rightarrow f}(x)\}_{x \in \mathcal{X}}$. Hence, without loss of generality, we can assume $\sum_x m_{f \rightarrow v}(x) = \sum_x m_{v \rightarrow f}(x) = 1$. Since the first derivatives are

$$\begin{aligned}\frac{\partial F(\{m_{v \rightarrow f}\}, \{m_{f \rightarrow v}\})}{\partial m_{f \rightarrow v}(x)} &= l \frac{m_{f \rightarrow v}(x)^{l-1}}{Z_v} - l \frac{m_{v \rightarrow f}(x)}{Z_{fv}} \\ \frac{\partial F(\{m_{v \rightarrow f}\}, \{m_{f \rightarrow v}\})}{\partial m_{v \rightarrow f}(x)} &= \frac{l}{r} \frac{\sum_{i=1}^r \sum_{\mathbf{x} \setminus x_i} f(\mathbf{x}) \prod_{j \neq i} m_{v \rightarrow f}(x_j)}{Z_f} \\ &\quad - l \frac{m_{f \rightarrow v}(x)}{Z_{fv}}\end{aligned}$$

the saddle point condition is

$$\begin{aligned}m_{v \rightarrow f}(x) &= \frac{Z_{fv}}{Z_v} m_{f \rightarrow v}(x)^{l-1} \\ m_{f \rightarrow v}(x) &= \frac{1}{r} \frac{Z_{fv}}{Z_f} \sum_{i=1}^r \sum_{\mathbf{x} \setminus x_i} f(\mathbf{x}) \prod_{j=1, j \neq i}^r m_{v \rightarrow f}(x_j).\end{aligned}$$

Although these points take minimal on any lines parallel to axis of $\{m_{f \rightarrow v}(x)\}$ or $\{m_{v \rightarrow f}(x)\}$, these points are not necessarily minimal.

APPENDIX C
MAGNETIC FIELD MODEL

Although we have only considered the random regular factor graph ensembles, the method can be generalized straightforwardly to many ensembles. As a simple example, we introduce the random regular factor graph with magnetic field.

$$p(\mathbf{x}) = \frac{1}{Z} \prod_a f(x_{\partial a}) \prod_i h(x_i).$$

Here, there are degree one factor nodes for each variable node. In the same way, the annealed free energy can be calculated.

$$\begin{aligned}\lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}[Z(\{v\}, \{\mu\})] \\ = \frac{l}{r} \mathcal{H}(\{\mu\}) - (l-1) \mathcal{H}(\{v\}) + \frac{l}{r} \sum_{\mathbf{x} \in \mathcal{X}^r} \mu(\mathbf{x}) \log f(\mathbf{x}) \\ + \sum_x v(x) \log h(x).\end{aligned}$$

Lemma 9.

$$\begin{aligned}\lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}[Z] \\ = \max_{(m_{f \rightarrow v}(x), m_{v \rightarrow f}(x)) \in \mathcal{S}} \left\{ \frac{l}{r} \log Z_f + \log Z_v - l \log Z_{fv} \right\}\end{aligned}$$

where \mathcal{S} denotes the set of saddle points of the function in max, and where

$$\begin{aligned} Z_v &:= \sum_x h(x) m_{f \rightarrow v}(x)^l \\ Z_f &:= \sum_{\mathbf{x}} f(\mathbf{x}) \prod_{i=1}^r m_{v \rightarrow f}(x_i) \\ Z_{fv} &:= \sum_x m_{f \rightarrow v}(x) m_{v \rightarrow f}(x) \end{aligned}$$

The stationary condition is

$$\begin{aligned} m_{v \rightarrow f}(x) &\propto h(x) m_{f \rightarrow v}(x)^{l-1} \\ m_{f \rightarrow v}(x) &\propto \sum_{i=1}^r \sum_{\substack{\mathbf{x} \setminus x_i \\ x_i = x}} f(\mathbf{x}) \prod_{j=1, j \neq i}^r m_{v \rightarrow f}(x_j). \end{aligned}$$

The above stationary condition is related to the stationary condition of maximization of the contribution $Z(v, \mu)$ with fixed variable type v in Section IV.

APPENDIX D PROOF OF LEMMA 4

Generally, when we have additional linear constraints

$$\begin{aligned} \sum_{x \in \mathcal{X}} a_k(x) v(x) &= b_k, \quad \text{for } k = 1, 2, \dots, s \\ \sum_{\mathbf{x}' \in \mathcal{X}} c_k(\mathbf{x}') \mu(\mathbf{x}') &= d_k, \quad \text{for } k = 1, 2, \dots, t \end{aligned}$$

in the maximization problem of $-F_{\text{Bethe}}(\{v\}, \{\mu\})$, the stationary condition is

$$\begin{aligned} \mu(\mathbf{x}) &\propto \left(\prod_{k=1}^t g_k^{c_k(\mathbf{x})} \right) f(\mathbf{x}) \prod_{i=1}^r m_{v \rightarrow f}(x_i) \\ v(x) &\propto \left(\prod_{k=1}^s h_k^{a_k(x)} \right) m_{f \rightarrow v}(x)^l \\ m_{v \rightarrow f}(x) &\propto \left(\prod_{k=1}^s h_k^{a_k(x)} \right) m_{f \rightarrow v}(x)^{l-1} \\ m_{f \rightarrow v}(x) &\propto \sum_{i=1}^r \sum_{\substack{\mathbf{x} \setminus x_i \\ x_i = x}} \left(\prod_{k=1}^t g_k^{c_k(\mathbf{x})} \right) f(\mathbf{x}) \prod_{j=1, j \neq i}^r m_{v \rightarrow f}(x_j). \end{aligned}$$

where $\{h_k \geq 0\}_{k=1, \dots, s}$ and $\{g_k \geq 0\}_{k=1, \dots, t}$ are auxiliary variables.

Proof:

$$\begin{aligned} L(v, \mu; \lambda, \rho, \eta, \zeta, \tau) &= -F_{\text{Bethe}} \\ &+ \lambda \left(\sum_{x \in \mathcal{X}} v(x) - 1 \right) + \sum_{k=1}^s \eta_k \left(\sum_{x \in \mathcal{X}} a_k(x) v(x) - b_k \right) \\ &+ \frac{l}{r} \rho \left(\sum_{\mathbf{x} \in \mathcal{X}^r} \mu(\mathbf{x}) - 1 \right) + \frac{l}{r} \sum_{k=1}^t \zeta_k \left(\sum_{\mathbf{x} \in \mathcal{X}^r} c_k(\mathbf{x}) \mu(\mathbf{x}) - d_k \right) \\ &+ \sum_{z \in \mathcal{X}} \tau(z) \left(\frac{l}{r} \sum_{i=1}^r \sum_{\substack{\mathbf{x} \setminus x_i \\ x_i = z}} \mu(\mathbf{x}) - l v(z) \right). \end{aligned}$$

$$\begin{aligned} \frac{\partial L}{\partial v(x)} &= (l-1)(\log v(x) + 1) + \lambda + \sum_{k=1}^s \eta_k a_k(x) - l \tau(x) \\ \frac{\partial L}{\partial \mu(\mathbf{x})} &= -\frac{l}{r} (\log \mu(\mathbf{x}) + 1) + \frac{l}{r} \log f(\mathbf{x}) + \frac{l}{r} \rho \\ &+ \frac{l}{r} \sum_{k=1}^t \zeta_k c_k(\mathbf{x}) + \frac{l}{r} \sum_{i=1}^r \tau(x_i). \end{aligned}$$

Let

$$\begin{aligned} \tau(x) &:= \log m_{v \rightarrow f}(x) =: \log \left(\frac{1}{Z_{v \rightarrow f}} \left(\prod_{k=1}^s h_k^{a_k(x)} \right) m_{f \rightarrow v}(x)^{l-1} \right) \\ \eta_k &:= \log h_k \\ \zeta_k &:= \log g_k. \end{aligned}$$

The rest of the proof is same as the proof of Lemma 1. \blacksquare

APPENDIX E PROOF OF LEMMA 7

We use the following relation.

$$\lim_{n \rightarrow 0} \frac{1}{n} \log \langle A^n \rangle = \langle \log A \rangle$$

where A is a random variable and $\langle \cdot \rangle$ denotes an expectation.

$$\begin{aligned} Z_v &= \sum_{\mathbf{x} \in \mathcal{X}^n} \left(\int \prod_{i=1}^n M_{v \rightarrow f}(x_i) dP \right)^l \\ &= \int \cdots \int \left(\prod_{j=1}^l dP_j \right) \left(\sum_x \prod_{j=1}^l M_{v \rightarrow f}^{(j)}(x) \right)^n \end{aligned}$$

Hence,

$$\lim_{n \rightarrow 0} \frac{1}{n} \log Z_v = \langle \log Z_v \rangle$$

The derivation of Z_f and Z_{fv} are similar.

The derivation of the saddle point equations are omitted since it is straightforward.

APPENDIX F REGULAR LDPC CODES

Corollary 10 ((Litsyn and Shevelev, 2002), (Burshtein and Miller, 2004)). *Growth rate of (l, r) -regular LDPC code ensemble is*

$$\begin{aligned} G(\omega) &= \frac{l}{r} \log \frac{1 + z''}{2} \\ &+ \log \left[e^h \left(\frac{1 + y'}{2} \right)^l + e^{-h} \left(\frac{1 - y'}{2} \right)^l \right] \\ &\quad - l \log \frac{1 + y' z'}{2} - \omega' h \end{aligned}$$

where $\omega' := 1 - 2\omega$ and

$$\begin{aligned} \omega' &= \tanh(h + l \tanh^{-1}(y')) \\ y' &= z'^{r-1} \\ z' &= \tanh(h + (l-1) \tanh^{-1}(y')). \end{aligned}$$

This result can be easily understood from Lemma 4 and 5 subject to by observing the following correspondings,

$$\begin{aligned}\omega' &= v(0) - v(1), & h &= (-1)^x \log h(x) \\ z' &= m_{v \rightarrow f}(0) - m_{v \rightarrow f}(1), & y' &= m_{f \rightarrow v}(0) - m_{f \rightarrow v}(1)\end{aligned}$$

and

$$\begin{aligned}Z_f &= \log \frac{1+z'^r}{2} \\ Z_v &= \log \left[e^h \left(\frac{1+y'}{2} \right)^l + e^{-h} \left(\frac{1-y'}{2} \right)^l \right] \\ Z_{fv} &= \log \frac{1+y'z'}{2} \\ \sum_x v(x) \log h(x) &= \omega' h.\end{aligned}$$

This result is also obtained by using the combinatorial method in [3] and change of variables [4]

$$h = -\frac{1}{2} \log x, \quad y' = \frac{1-y}{1+y}, \quad z' = \frac{1-z}{1+z}.$$

But the proof of this paper is much more meaningful.

APPENDIX G RANDOM MAGNETIC FIELD MODEL

In this appendix, we consider the random magnetic field model.

$$\begin{aligned}p(\mathbf{x} | \{h_i\}) &= \frac{1}{Z(\{h_i\})} \prod_a f(\mathbf{x}_{\partial a}) \prod_i h_i(x_i) \\ Z(\{h_i\}) &= \sum_{\mathbf{x}} \prod_a f(\mathbf{x}_{\partial a}) \prod_i h_i(x_i).\end{aligned}$$

Here, $\{h_i\}$ independently and identically distributed according to the distribution $P_H(h)$ on a finite set \mathcal{H} of nonnegative function on \mathcal{X} . In statistical physics, $h_i(x_i)$ represents random magnetic field. As a posteriori probability of LDPC codes, h_i corresponds to output of a channel. We now consider $\lim_{N \rightarrow \infty} 1/N \mathbb{E}_{\{h_i\}} [\log \mathbb{E}[Z(\{h_i\})]]$. Since $\mathbb{E}[Z(\{h_i\})]$ depends on $\{h_i\}$ only through the type of $\{h_i\}$, and since $1/N \log \mathbb{E}[Z(\{h_i\})] = O(1)$ for any $\{h_i\}$, we only have to deal with typical $\{h_i\}$. Let $v(x, h)$ denotes the number of variable nodes of value x and whose corresponding factor is h . The factor-type $u(\mathbf{x}, \mathbf{h})$ is defined in the same way. Then, it holds

$$Z = \sum_{v, u} N(v, u) \prod_{(\mathbf{x}, \mathbf{h}) \in \mathcal{X}^r \times \mathcal{H}^r} f(\mathbf{x})^{u(\mathbf{x}, \mathbf{h})} \prod_{(x, h) \in \mathcal{X} \times \mathcal{H}} h(x)^{v(x, h)}$$

For typical $\{h_i\}$, it holds

$$\mathbb{E}[N(v, u)] = \prod_{h \in \mathcal{H}} \binom{NP_H(h)}{\{v(x, h)\}_{x \in \mathcal{X}}} \binom{\frac{l}{r}N}{\{u(\mathbf{x}, \mathbf{h})\}_{(\mathbf{x}, \mathbf{h}) \in \mathcal{X}^r \times \mathcal{H}^r}} \cdot \frac{\prod_{(x, h) \in \mathcal{X} \times \mathcal{H}} (v(x, h)l)!}{(Nl)!}.$$

Hence, the problem is maximization of

$$\begin{aligned}& \frac{l}{r} \mathcal{H}(\mu) - (l-1) \mathcal{H}(v) - \mathcal{H}(P_H) \\ & + \frac{l}{r} \sum_{(\mathbf{x}, \mathbf{h}) \in \mathcal{X}^r \times \mathcal{H}^r} \mu(\mathbf{x}, \mathbf{h}) \log f(\mathbf{x}) + \sum_{(x, h) \in \mathcal{X} \times \mathcal{H}} v(x, h) \log h(x)\end{aligned}$$

$$\begin{aligned}v(x, h) &\geq 0, & \mu(\mathbf{x}, \mathbf{h}) &\geq 0 \\ \sum_x v(x, h) &= P_H(h), & \sum_{\mathbf{x}, \mathbf{h}} \mu(\mathbf{x}, \mathbf{h}) &= 1\end{aligned}$$

$$\frac{1}{r} \sum_{i=1}^r \sum_{\substack{(\mathbf{x}, \mathbf{h}) \setminus (x_i, h_i) \\ x_i=z, h_i=h}} \mu(\mathbf{x}, \mathbf{h}) = v(z, h), \forall z \in \mathcal{X}, h \in \mathcal{H}.$$

Lemma 11. *The stationary conditions are*

$$\begin{aligned}\mu(\mathbf{x}, \mathbf{h}) &\propto f(\mathbf{x}) \prod_{i=1}^r m_{v \rightarrow f}(x_i, h_i) \\ v(x, h) &\propto g(h) h(x) m_{f \rightarrow v}(x)^l\end{aligned}$$

where

$$P_H(h) \propto g(h) \sum_{x \in \mathcal{X}} h(x) m_{f \rightarrow v}(x)^l \quad (16)$$

$$\begin{aligned}m_{v \rightarrow f}(x, h) &\propto g(h) h(x) m_{f \rightarrow v}(x)^{l-1} \\ m_{v \rightarrow f}(x) &\propto \sum_h m_{v \rightarrow f}(x, h)\end{aligned} \quad (17)$$

$$m_{f \rightarrow v}(x) \propto \sum_{i=1}^r \sum_{\substack{\mathbf{x} \setminus x_i \\ x_i=x}} f(\mathbf{x}) \prod_{j=1, j \neq i}^r m_{v \rightarrow f}(x_j). \quad (18)$$

Here $m_{v \rightarrow f}(x)$, $m_{f \rightarrow v}(x)$ and $g(h)$ are auxiliary functions satisfying $\sum_{x \in \mathcal{X}} m_{v \rightarrow f}(x) = \sum_{x \in \mathcal{X}} m_{f \rightarrow v}(x) = 1$, and $g(h) \geq 0$.

Lemma 12.

$$\begin{aligned}& \lim_{N \rightarrow \infty} \mathbb{E}_{\{h_i\}} [\log \mathbb{E}[Z]] \\ &= \max_{(m_{f \rightarrow v}(x), m_{v \rightarrow f}(x), g(h)) \in \mathcal{S}} \left\{ \frac{l}{r} \log Z_f + \log Z_v - l \log Z_{fv} \right. \\ & \quad \left. + \sum_h P_H(h) \log \frac{P_H(h)}{g(h)} \right\} \\ &= \max_{(m_{f \rightarrow v}(x), m_{v \rightarrow f}(x)) \in \mathcal{S}} \left\{ \frac{l}{r} \log Z_f + \sum_h P_H(h) \log Z_v(h) \right. \\ & \quad \left. - l \log Z_{fv} \right\}\end{aligned}$$

where \mathcal{S} denotes the set of saddle points of the function in max, and where

$$\begin{aligned}Z_f &:= \sum_{\mathbf{x} \in \mathcal{X}^r} f(\mathbf{x}) \prod_{i=1}^r m_{v \rightarrow f}(x_i) \\ Z_v(h) &:= \sum_{x \in \mathcal{X}} h(x) m_{f \rightarrow v}(x)^l \\ Z_v &:= \sum_{h \in \mathcal{H}} g(h) Z_v(h) \\ Z_{fv} &:= \sum_{x \in \mathcal{X}} m_{f \rightarrow v}(x) m_{v \rightarrow f}(x).\end{aligned}$$

The conditions of saddle point are (16) to (18).

APPENDIX H
IRREGULAR AND POISSON ENSEMBLES

A. Irregular ensemble

The result can be generalized for irregular ensembles. Let \mathcal{D}_v and \mathcal{D}_c denote the set of degrees of variable nodes and check nodes, respectively. Let L_i and R_j denote the degree distribution of variable nodes and check nodes from node perspective for $i \in \mathcal{D}_v$ and $j \in \mathcal{D}_c$, respectively. Assume the factor corresponding to degree j factor nodes is $f_j(\mathbf{x})$ for $j \in \mathcal{D}_c$. Let $v(i, x)$ denotes the number of variable nodes of degree i and value x . The factor-type $u(j, \mathbf{i}, \mathbf{x})$ is defined in the same way.

$$\mathbb{E}[N(v, u)] = \prod_{i \in \mathcal{D}_v} \left(\frac{NL_i}{\{v(i, x)\}_{x \in \mathcal{X}}} \right) \times \prod_{j \in \mathcal{D}_c} \left(\frac{\frac{L'(1)}{R'(1)} NR_j}{\{u(j, \mathbf{i}, \mathbf{x})\}_{(\mathbf{i}, \mathbf{x}) \in \mathcal{D}_v^j \times \mathcal{X}^j}} \right) \frac{\prod_{(i, x) \in \mathcal{D}_v \times \mathcal{X}} (v(i, x) i)!}{(NL'(1))!}$$

The problem is the maximization of

$$\begin{aligned} & \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}[Z(v, \mu)] \\ &= \frac{L'(1)}{R'(1)} \sum_{j \in \mathcal{D}_c} R_j \mathcal{H}(\mu_j) - \sum_{i \in \mathcal{D}_v} L_i (i-1) \mathcal{H}(v_i) - L'(1) \mathcal{H}(L_i i) \\ &+ \frac{L'(1)}{R'(1)} \sum_{j \in \mathcal{D}_c} \sum_{(\mathbf{i}, \mathbf{x}) \in \mathcal{D}_v^j \times \mathcal{X}^j} \mu(j, \mathbf{i}, \mathbf{x}) \log f_j(\mathbf{x}) \end{aligned}$$

subject to

$$\begin{aligned} v(i, x) &\geq 0, & \mu(j, \mathbf{i}, \mathbf{x}) &\geq 0 \\ \sum_{x \in \mathcal{X}} v(i, x) &= L_i, & \sum_{(\mathbf{i}, \mathbf{x}) \in \mathcal{D}_v^j \times \mathcal{X}^j} \mu(j, \mathbf{i}, \mathbf{x}) &= R_j \end{aligned}$$

$$\frac{L'(1)}{R'(1)} \sum_{j \in \mathcal{D}_c} \sum_{k=1}^j \sum_{(\mathbf{i}, \mathbf{x}), (i_k, x_k) = (i, x)} \mu(j, \mathbf{i}, \mathbf{x}) = i v(i, x).$$

We obtain the following stationary conditions.

$$\begin{aligned} \mu(j, \mathbf{i}, \mathbf{x}) &\propto r(j) f_j(\mathbf{x}) \prod_{k=1}^j m_{v \rightarrow f}(i_k, x_k) \\ v(i, x) &\propto l(i) m_{f \rightarrow v}(x)^i \\ L_i &\propto l(i) \sum_{x \in \mathcal{X}} m_{f \rightarrow v}(x)^i \end{aligned} \quad (19)$$

$$R_j \propto r(j) \sum_{\mathbf{x} \in \mathcal{X}^j} f_j(\mathbf{x}) \prod_{k=1}^j m_{v \rightarrow f}(x_k) \quad (20)$$

$$\begin{aligned} m_{v \rightarrow f}(i, x) &\propto i l(i) m_{f \rightarrow v}(x)^{i-1} \\ m_{v \rightarrow f}(x) &\propto \sum_{i \in \mathcal{D}_v} i l(i) m_{f \rightarrow v}(x)^{i-1} \end{aligned} \quad (21)$$

$$m_{f \rightarrow v}(x) \propto \sum_{j \in \mathcal{D}_c} \sum_{t=1}^j \sum_{\substack{\mathbf{x} \in \mathcal{X}^j \\ x_t = x}} r(j) f_j(\mathbf{x}) \prod_{k=1, k \neq t}^j m_{v \rightarrow f}(x_k) \quad (22)$$

Lemma 13.

$$\begin{aligned} & \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}[Z] \\ &= \max_{(m_{v \rightarrow f}(x), m_{f \rightarrow v}(x), l(i), r(j)) \in \mathcal{S}} \left\{ \frac{L'(1)}{R'(1)} \log Z_f + \log Z_v \right. \\ &\quad \left. - L'(1) \log Z_{fv} + \frac{L'(1)}{R'(1)} \sum_{j \in \mathcal{D}_c} R_j \log \frac{R_j}{r(j)} + \sum_{i \in \mathcal{D}_v} L_i \log \frac{L_i}{l(i)} \right\} \\ &= \max_{(m_{v \rightarrow f}(x), m_{f \rightarrow v}(x)) \in \mathcal{S}} \left\{ \frac{L'(1)}{R'(1)} \sum_{j \in \mathcal{D}_c} \log Z_f(j) + \sum_{i \in \mathcal{D}_v} L_i \log Z_v(i) \right. \\ &\quad \left. - L'(1) \log Z_{fv} \right\} \end{aligned} \quad (23)$$

where \mathcal{S} denotes the set of saddle points of the function in max, and where

$$\begin{aligned} Z_f(j) &:= \sum_{\mathbf{x} \in \mathcal{X}^j} f_j(\mathbf{x}) \prod_{k=1}^j m_{v \rightarrow f}(x_k) \\ Z_v(i) &:= \sum_{x \in \mathcal{X}} m_{f \rightarrow v}(x)^i \\ Z_v &:= \sum_{i \in \mathcal{D}_v} l(i) Z_v(i) \\ Z_f &:= \sum_{j \in \mathcal{D}_c} r(j) Z_f(j) \\ Z_{fv} &:= \sum_{x \in \mathcal{X}} m_{f \rightarrow v}(x) m_{v \rightarrow f}(x) \end{aligned}$$

The stationary conditions are (19) to (22).

This second expression (23) is equivalent to the equations in [4].

B. Poisson ensemble

In this subsection, we deal with Poisson ensemble. There are N variable nodes and αN factor nodes. The degree of factor node is k . For each factor node, connecting variable nodes are chosen independently and uniformly from $N(N-1) \cdots (N-(k-1))$ ways. In the same way as other ensembles, we obtain

$$\begin{aligned} \mathbb{E}[N(v, u)] &= \binom{N}{\{v(x)\}_{x \in \mathcal{X}}} \binom{\alpha N}{\{u(\mathbf{x})\}_{\mathbf{x} \in \mathcal{X}^k}} \\ &\times \prod_{\mathbf{x} \in \mathcal{X}^k} \left(\frac{\prod_{x \in \mathcal{X}} v(x) (v(x)-1) \cdots (v(x) - (N_x(\mathbf{x}) - 1))}{N(N-1) \cdots (N-(k-1))} \right)^{u(\mathbf{x})} \end{aligned}$$

where $N_x(\mathbf{x})$ denotes the number of x in \mathbf{x} . The problem is maximization of

$$\begin{aligned} & \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}[Z(v, \mu)] \\ &= \alpha \mathcal{H}(\mu) + \mathcal{H}(v) + \alpha \sum_{\mathbf{x} \in \mathcal{X}^k} \mu(\mathbf{x}) \log \left(\prod_{i=1}^k v(x_i) \right) \\ &\quad + \alpha \sum_{\mathbf{x} \in \mathcal{X}^k} \mu(\mathbf{x}) \log f(\mathbf{x}) \\ &= -\alpha \mathcal{D}(\mu \| v^k) + \mathcal{H}(v) + \alpha \sum_{\mathbf{x} \in \mathcal{X}^k} \mu(\mathbf{x}) \log f(\mathbf{x}) \end{aligned}$$

subject to

$$\begin{aligned} v(x) &\geq 0, & \mu(\mathbf{x}) &\geq 0 \\ \sum_x v(x) &= 1, & \sum_{\mathbf{x}} \mu(\mathbf{x}) &= 1. \end{aligned}$$

This is also similar to the minimization of Bethe free energy since (1) is also written as

$$\begin{aligned} \mathcal{H}_{\text{Bethe}}(b_i, b_a) &= - \sum_a \sum_{\mathbf{x}_{\partial a}} b_a(\mathbf{x}_{\partial a}) \log \frac{b_a(\mathbf{x}_{\partial a})}{\prod_{j \in \partial a} b_j(x_j)} \\ &\quad - \sum_i \sum_{x_i} b_i(x_i) \log b_i(x_i). \end{aligned}$$

The derivation of the following lemma is omitted for lack of space.

Lemma 14.

$$\begin{aligned} &\lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}[Z(v, \mu)] \\ &= \max_{(m_{f \rightarrow v}(x), m_{v \rightarrow f}(x), e) \in \mathcal{S}} \left\{ \alpha \log Z_f + \log Z_v \right. \\ &\quad \left. - e \sum_x m_{v \rightarrow f}(x) m_{f \rightarrow v}(x) \right\} \quad (24) \end{aligned}$$

where \mathcal{S} denotes the set of saddle points of the function in max, and where

$$\begin{aligned} Z_f &:= \sum_{\mathbf{x}} f(\mathbf{x}) \prod_{i=1}^k m_{v \rightarrow f}(x_i) \\ Z_v &:= \sum_x \exp\{e m_{f \rightarrow v}(x)\} \end{aligned}$$

The conditions of saddle point are

$$\begin{aligned} m_{f \rightarrow v}(x) &= \frac{\alpha}{e Z_f} \sum_{i=1}^k \sum_{\substack{\mathbf{x} \in \mathcal{X}^k \\ x_i = x}} f(\mathbf{x}) \prod_{j=1, j \neq i}^k m_{v \rightarrow f}(x_j) \\ m_{v \rightarrow f}(x) &\propto \exp\{e m_{f \rightarrow v}(x)\} \\ &= 1 + e m_{f \rightarrow v}(x) + \frac{(e m_{f \rightarrow v}(x))^2}{2!} + \dots \end{aligned}$$

Here, e can be regarded as mean of Poisson distribution expressing the degree distribution of variable nodes. Note that the third term of (24) evaluated at saddle points is αk .

APPENDIX I

STABILITY OF THE PARAMAGNETIC SOLUTION

Assume $\sum_{i=1}^r \sum_{\substack{\mathbf{x} \in \mathcal{X}^r \\ x_i = x}} f(\mathbf{x})$ is constant among all $x \in \mathcal{X}$. Let $v(x) = m_{f \rightarrow v}^P(x) = m_{v \rightarrow f}^P(x) = 1/q$ for all $x \in \mathcal{X}$. Let us start the algorithm in Section IV from

$$m_{f \rightarrow v}(x) \propto m_{f \rightarrow v}^P(x) + \delta(x).$$

By linear approximation,

$$\begin{aligned} m_{v \rightarrow f}^+(x) &\propto \frac{v(x)}{m_{f \rightarrow v}^P(x) + \delta(x)} \\ &= \frac{v(x)}{m_{f \rightarrow v}^P(x)} \left[1 - \frac{\delta(x)}{m_{f \rightarrow v}^P(x)} + \Theta(\delta(x)^2) \right] \\ &= 1 - q\delta(x) + \Theta(\delta(x)^2) \\ m_{f \rightarrow v}^+(x) &\propto \sum_{i=1}^r \sum_{\substack{\mathbf{x} \in \mathcal{X}^r \\ x_i = x}} f(\mathbf{x}) \prod_{j=1, j \neq i}^r m_{v \rightarrow f}^+(x_j) \\ &\propto \sum_{i=1}^r \sum_{\substack{\mathbf{x} \in \mathcal{X}^r \\ x_i = x}} f(\mathbf{x}) \prod_{j=1, j \neq i}^r (1 - q\delta(x_j) + \Theta(\delta(x_j)^2)) \\ &= \sum_{i=1}^r \sum_{\substack{\mathbf{x} \in \mathcal{X}^r \\ x_i = x}} f(\mathbf{x}) - q \sum_{i=1}^r \sum_{\substack{\mathbf{x} \in \mathcal{X}^r \\ x_i = x}} f(\mathbf{x}) \left(\sum_{j=1, j \neq i}^r \delta(x_j) \right) + \sum_{x \in \mathcal{X}} \Theta(\delta(x)^2) \\ &\propto \frac{1}{q} - \frac{\sum_{i=1}^r \sum_{\substack{\mathbf{x} \in \mathcal{X}^r \\ x_i = x}} f(\mathbf{x}) \left(\sum_{j=1, j \neq i}^r \delta(x_j) \right)}{\sum_{i=1}^r \sum_{\substack{\mathbf{x} \in \mathcal{X}^r \\ x_i = x}} f(\mathbf{x})} + \sum_{x \in \mathcal{X}} \Theta(\delta(x)^2) \end{aligned}$$

Let

$$\delta^+(x) := - \frac{\sum_{i=1}^r \sum_{\substack{\mathbf{x} \in \mathcal{X}^r \\ x_i = x}} f(\mathbf{x}) \left(\sum_{j=1, j \neq i}^r \delta(x_j) \right)}{\sum_{i=1}^r \sum_{\substack{\mathbf{x} \in \mathcal{X}^r \\ x_i = x}} f(\mathbf{x})}.$$

We now consider the linear operator A defined by $A(\{\delta(x)\}_{x \in \mathcal{X}}) = \{\delta^+(x)\}_{x \in \mathcal{X}}$. The all 1 vector is an eigenvector of A with eigenvalue $-(r-1)$. The stability condition of the paramagnetic solution is that absolute values of eigenvalues of A not corresponding to the all 1 vector are smaller than 1. For the binary CSP (14), the matrix A is a symmetric 2×2 matrix where

$$\begin{aligned} A_{11} = A_{22} &= - \frac{(r-1) \sum_{i=0}^{\frac{r}{2}-k-1} \binom{r-1}{i} + \left(\frac{r-1}{2-k}\right) \left(\frac{r}{2} + k - 1\right)}{2 \sum_{i=0}^{\frac{r}{2}-k-1} \binom{r-1}{i} + \left(\frac{r-1}{2-k}\right)} \\ A_{12} = A_{21} &= - \frac{(r-1) \sum_{i=0}^{\frac{r}{2}-k-1} \binom{r-1}{i} + \left(\frac{r-1}{2-k}\right) \left(\frac{r}{2} - k\right)}{2 \sum_{i=0}^{\frac{r}{2}-k-1} \binom{r-1}{i} + \left(\frac{r-1}{2-k}\right)}. \end{aligned}$$

The eigenvalues of A are $A_{11} + A_{12}$ and $A_{11} - A_{12}$ whose eigenvectors are $[1 \ 1]^T$ and $[1 \ -1]^T$, respectively. We can easily confirm that

$$\begin{aligned} A_{11} + A_{12} &= -(r-1) \\ A_{11} - A_{12} &= - \frac{\left(\frac{r-1}{2-k}\right)(2k-1)}{2 \sum_{i=0}^{\frac{r}{2}-k-1} \binom{r-1}{i} + \left(\frac{r-1}{2-k}\right)}. \end{aligned}$$

Hence, the stability condition is

$$\frac{\left(\frac{r-1}{2-k}\right)(2k-1)}{2 \sum_{i=0}^{\frac{r}{2}-k-1} \binom{r-1}{i} + \left(\frac{r-1}{2-k}\right)} < 1.$$

For $r = 20$ and $k = 1, 2, 3$, the left-hand side of the condition is 0.23883, 0.859049 and 1.825917, respectively. This result is consistent with the numerical calculation result in Fig. 1.